# Image Fusion of Video Images and Geo-localization for UAV Applications

K.Senthil Kumar[1], Kavitha.G[2], Subramanian.R[3] and Marwan[4]
[1, 3, 4] Division of Avionics
[2] Department of Electronics and Communication Engineering,
Madras Institute of Technology,
Anna University Chennai, India
[1] ksk_mit@annauniv.edu
[2] kavithag_mit@annauniv.edu
[3] subramanian.r.88@gmail.com
[4] munna860@gmail.com

***Abstract*—We present in this paper a very fine method for determining the location of a ground based target when viewed from an Unmanned Aerial Vehicle (UAV). By determining the pixel coordinates on the video frame and by using a range finder the target's geo-location is determined in the North-East-Down (NED) frame. The contribution of this method is that the target can be localized to within 9m when view from an altitude of 2500m and down to 1m from an altitude of 100m. This method offers a highly versatile tracking and geo-localisation technique that has very good number of advantages over the previously suggested methods. Some of the key factors that differentiate our method from its predecessors are:**

1) **Day and night time operation**
2) **All weather operation**
3) **Highly accurate positioning of target in terms of latitude-longitude (GPS) and altitude.**
4) **Automatic gimbaled operation of the camera once target is locked**
5) **Tracking is possible even when the target stops moving**
6) **Independent of target (moving or stationary)**
7) **No terrain database is required**
8) **Instantaneous target geolocalisation is possible**

***Index Terms*—first term, second term, third term, fourth term, fifth term, sixth term**

## I. Introduction

Unmanned Aerial Vehicles (UAV) exists since they have the ability and advantage to perform dangerous tasks. UAV's have an advantage of not putting human lives at risk. Tasks such as tracking and reconnaissance require the UAV to determine the location of the target for military actions such as dropping GPS guided bombs. This method presents a method for determining the GPS coordinates of the target independent of whether the target is moving or stationary and independent of the type of target (vehicles, human or animals). The method uses different approaches to track the target using thermal imaging and to determine the targets by using an UAV fitted with a gimbal assembly of a RGB vision camera, thermal camera and a range finder. In this paper we have assumed the target is identified by a human at the ground control station (GCS). The UAV requires two persons for operation. One person operates the UAV (UAV pilot) while the other person operates the gimbal assembly (UAV operator). Once the UAV operator identifies the target the UAV system locks onto the target and the gimbal assembly automatically adjusts its azimuth and elevation to keep the target in the Field of View (FOV). This method uses three steps in the process:

### *1.1. Thermal and visual image fusion of airborne video*

The use of thermal imaging technology and the eventual capabilities has increased dramatically in this technological era. Thermal imaging used to be a very expensive technology for military users only. Today, more applications are involved with the help of thermal imaging. Sensor information fusion which involves the process of combining various sensing modalities gives a realistic view. The subset of the sensor information does not collectively reveal the necessary data. With the development of new imaging sensors arises the need of a meaningful combination of all employed imaging sources. Image fusion of visual and thermal sensing outputs adds a new dimension in making the target tracking application of UAV more reliable. Target tracking at instances of smoke, fog and cloudy conditions gets improved. Target identification, localisation, filtering and data association forms an important application of the fusion process. Thus an effective surveillance and reconnaissance system can be formed.

### 1.2. Tracking of relevant target

The fused thermal video gives a clear distinction of the target from its environment. The tracking module uses Horn-Schunck method of optical flow to determine motion. Previously used methods require the operator to control the gimbal assembly to keep the target in the Field of View (FOV). Using the current method the target is tracked and there is no need to define the target to the system once tracking commences.




DOI: 01.IJIT.01.02.33

### 1.3. *Target geo-localization*

Once the target is tracked on the pixel coordinate frame, data from the range finder is combined and then converted to the NED frame. Hence the target can be tracked automatically using a feedback control mechanism connected with the gimbal assembly and its instantaneous GPS coordinates can be determined in real time. Compared to previous methods which were had an accuracy of up to 8m, this method is capable of determining the position of the target with accuracy within 1m on the horizontal plane and around 2meters error in altitude. Hence the complete 3D coordinates of the target can be determined.

## II. PROPOSED METHOD

The proposed method can be split into the following subdivisions as:

### A. *Thermal and visual image fusion process*

Thermal images have a valuable advantage over the visual images. Thermal images do not depend on the illumination, the output is the projection of thermal sensors of the emissions of heat of the objects. This unique merit gives rise for effective segmentation of objects. Ultimately, surveillance measure using an UAV gets improved. Considering a visual and thermal video as output from the UAV to be obtained in the ground control station, the two videos are split into image. The visual and thermal image of the same frame is fused with by applying Haar wavelet transform to the fourth level. An inverse wavelet transform gets us the fused image. The considerations to be met here are same resolution images with same field of view.

### B. *Gimbal assembly*

The gimbal assembly constitutes of a thermal camera, a video camera having the same intrinsic parameters as the thermal camera and a range finder, all fitted onto the same gimbal so that they rotate altogether in the same direction in any axis. The gimbal assembly is positioned such that the thermal camera, video camera and the range finder are all pointing in the same direction or in other words that their line of sights are always parallel. The gimbal assembly is also fitted with accelerometers to measure the angle of elevation and azimuth of the optical sensors. The elevation and azimuth values are measured with respect to the body of the UAV.

The gimbal assembly has two functions:

1. Keeps the thermal and optic sensors aligned on the same platform.
2. Measures the azimuth (az) and elevation angle (el) and sends the values to the processor.

It should be noted that the gimbal attitude parameters are independent of the attitude of the UAV.

### C. *Tracking Process*

The tracking process uses the fused video result. Here the tracking algorithm uses Horn-Schunck method to track optical flow in the video. The Horn-Schunck method is superior to the Lucas-Kanade method because it is resistant to the 'aperture problem' and it is very suitable for UAV and airborne applications where the camera is mounted on a moving platform. Once the system starts tracking the target on the instructions given by the operator from the GCS, the centroid of the threshold area is computed for every frame. This centroid value is always maintained at the center of the image plane using servos controlling the gimbal assembly. The difference between the center of the image plane and the centroid value obtained while tracking is determined for the azimuth and elevation axes to make corrections and maintain the error as minimum in the two axes i.e., to provide feedback to the gimbal assembly to adjust itself so that the target is always near the center of the image plane. This ensures that there is no requirement for the operator at the GCS to control the gimbal assembly to focus on the target. The feedback mechanism takes care of this feature.

### D. *Geolocalization*

The gimbal assembly constitutes of a thermal camera, a video camera having the same intrinsic parameters as the thermal camera and a range finder, all fitted onto the same gimbal so that they rotate altogether in the same direction in any axis. The gimbal assembly is positioned such that the thermal camera, video camera and the range finder are all pointing in the same direction or in other words that their line of sights are always parallel. The gimbal assembly is also fitted with accelerometers to measure the angle of elevation and azimuth of the optical sensors. The elevation and azimuth values are measured with respect to the body of the UAV.

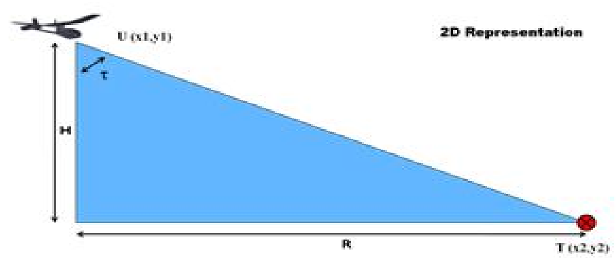


Fig. 1. A two dimensional representation of localizing the target

In Fig. 1 it can be illustrated that by knowing the value of Height of the UAV 'H' and angle of elevation of the gimbal assembly '$\tau$', the range R of the target on the horizontal plane from the UAV's can be determined with ease. It should be observed that in the above case that the target is considered to be at zero altitude. However if the target is on a different terrain such as a trench or a hill, then the above determination of the target's coordinates will yield wrong. As shown in Fig. 2, the actual target is at position $T(x_2,y_2)$. But by using the above method we will get another set of coordinates i.e., $T(x_2',y2')$, which are wrong.

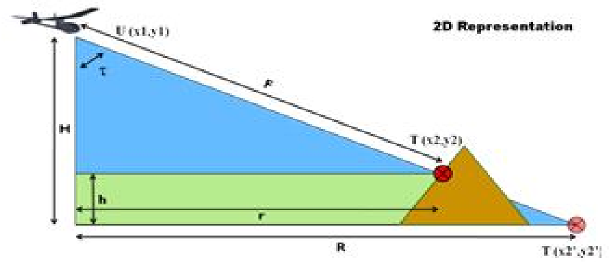


Fig. 2. A two dimensional representation of localizing the target on a hilly terrain




DOI: 01.IJIT.01.02.33

Similar work can be done in 3D, the only difference being that the angular value $\tau$ will be replaced by 'el' and 'az'. The Range Finder serves its purpose to overcome this problem. The range finder which is with the gimbal assembly provides us with the slant range 'F'. Now this value 'F' can be used to determine the ground distance 'R' between the UAV and the target. The three dimensional representation of the schematics are shown in Fig. 3 and Fig. 4.

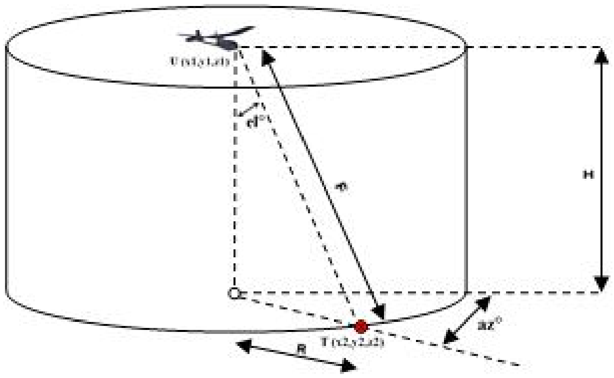


Fig. 3. A three dimensional representation of localizing the target

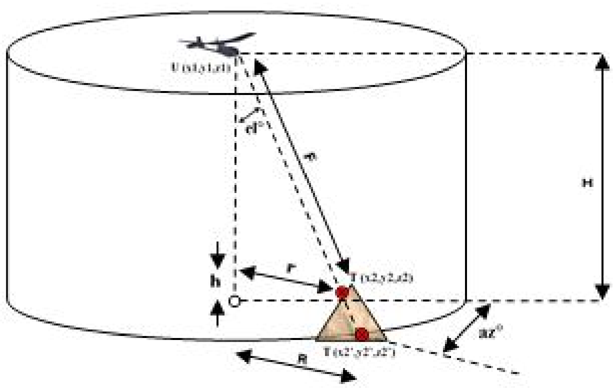


Fig. 4. A three dimensional representation of localizing the target on a hilly terrain

## II. RESULTS

The fusion, tracking and localization results are given below:

### *A. Thermal and Visual Image Fusion*

Image fusion results based on wavelet transform are discussed. Two images (Visual and thermal) are taken as shown in Fig. 5 and Fig. 6 and the fused image is obtained as an output as shown in Fig. 7. The visual image gives a realistic human sensing view. The thermal image identifies the target with the temperature difference coming into the picture with objects possessing different emissivity values. The fused image is obtained as a result of four level wavelet transform and the combined information is obtained. The fused image combines the complementary data. Image fusion results are with respect to the images taken from imagefusion.org image database collection.

Fig. 5. Visual image from the database collection

Fig. 6. Thermal image from the database collection

Fig. 7. Thermal and visual fused image

### *B. Geo-Localization*

The videos used for simulating tracking are self animated videos which help to understand the problems involved by introducing the unwanted components like gradient noise into the animation and hence the best and optimized method can be adopted. For the airborne video the Horn Schunk's method proves effective and the tracking results of the algorithm are. The following figures show the results obtained in the tracking section.




DOI: 01.IJIT.01.02.33

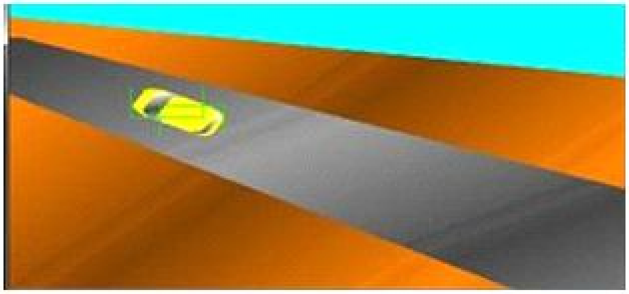

Fig. 8. Sample target tracking video

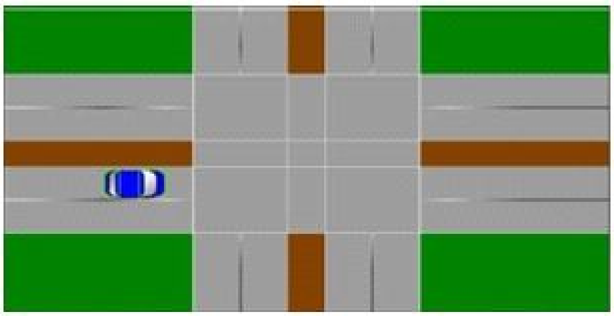

Fig. 9. Sample target tracking video

The centroid of the bounding box gives the coordinates (x,y) of the detected target on the image plane. This can be used only if the car is in motion. If the car stops then the centroid value returns to (0,0). This is shown graphically in Fig. 10.

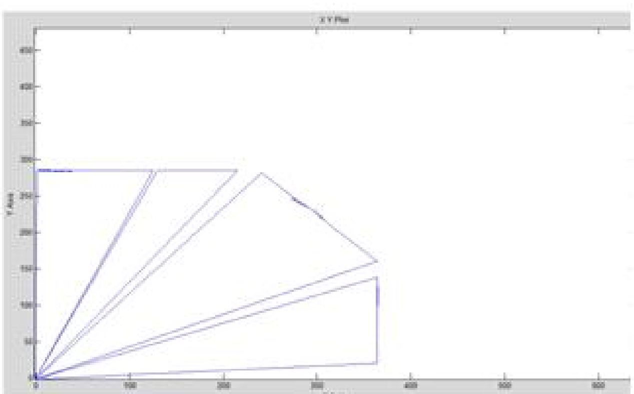

Fig. 10. Two Dimensional representation of the car's movement on a plane

However by storing the previous value of the target in a buffer it is possible to hold the position of the target when the target is not moving in the image plane. This is possible using a user customized position hold block as shown below.

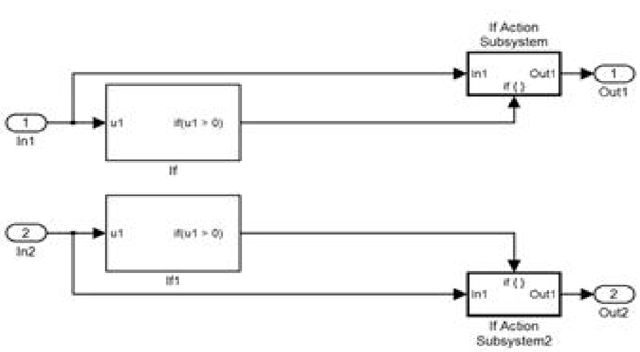


Fig. 11. IF Loop for Position Hold Function Block

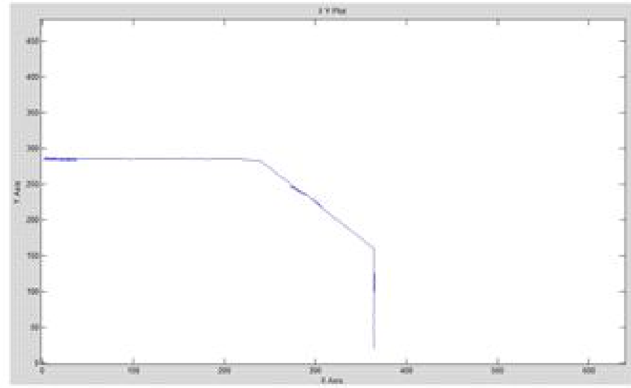

Fig. 12. Two Dimensional representation of the car's movement on a plane without the Position Hold Block

Results from MATLAB showing accuracy ranges after introduction of errors: The results obtained using the proposed method was found satisfactory even after introducing the errors which were previously discussed. Table I shows the tested accuracies at different variations in GPS altitude, gimbal assembly and range finder. The results in Table I show that the maximum and minimum accuracy of the method is 0.95 m and 21.8 m respectively.

TABLE. I
GEOLOCALISATION ACCURACY RESULTS

| | | |
|---|---|---|
| ***Minimum Accuracy*** | ***21.82087018*** | |
| ***Maximum Accuracy*** | ***0.948123798*** | |
| | | |
| Accuracy | Min | Max |
| Accuracy at H=2400 m | 0.948123798 | 21.820870 |
| Accuracy at H= 100 m | 0.948123798 | 21.820870 |
| Accuracy at F=2500 m | 8.314817257 | 21.820870 |
| Accuracy at F= 100 m | 0.948123798 | 1.2387153 |
| Accuracy at T= 5° | 0.948123798 | 21.820870 |
| Accuracy at T= 70° | 1.237528522 | 8.4877510 |

Fig. 13 depicts the accuracy of the target in graphical form. The green dot depicts the true position of the target and the red and blue dots depict the accuracy obtained by introducing positive and negative errors respectively. These values of accuracies were computed for different error values at varying altitude 'H', LRF range 'F' and gimbal angle 'τ'.

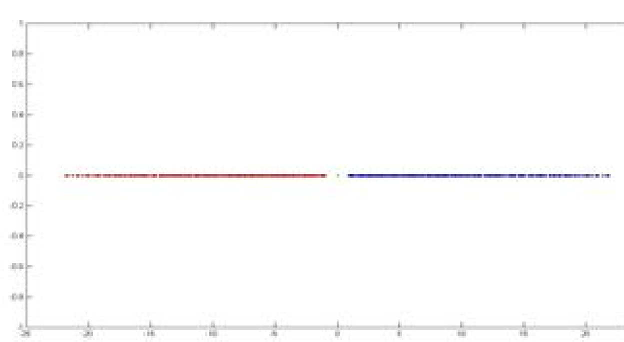

Fig. 13. Accuracy Distribution with respect to the true position




DOI: 01.IJIT.01.02. 33

## Conclusions

The airborne images obtained from an UAV are analyzed in ground control station. By using the thermal images, all weather and night operation are possible. Visual and thermal image fusion is done and the fused image is given for target tracking. This system has the benefit of enhanced target tracking application wherein only visual or thermal target tracking would not provide sufficient efficiency. Thus the image fusion process augments information leading to an improved system as a whole. The overall system incorporates segmentation, fusion and target tracking principles. By using the tracked coordinates on the image frame, the coordinates can be shifted to the NED frame which will give the GPS coordinates of the target. The method proves it robust application for the military and can prove efficient since it has an accuracy of 0.9m from an altitude of 100m.

## Acknowledgment

The authors wish to thank A, B, C. This work was supported in part by a grant from XYZ.

## References

[1] Y.Chena and C.Han, "Night-time Pedestrian Detection by Visual-Infrared Video Fusion," *Proceedings of 7th World congress on Intelligent Control and Automation*, China, 2008.

[2] Alex Leykin, Yang Ran and Riad Hammoud, "Thermal-Visible Video Fusion for Moving Target Tracking and Pedestrian Classification", *IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, 2007.

[3] Zhang Jin-Yu, Chen Yan and Huang Xian-Xiang, "IR Thermal Image Segmentation Based on Enhanced Genetic Algorithms and Two-Dimensional Classes Square Error", *Second International Conference on Information and Computing Science*, 2009.

[4] Daniel Olmeda, Arturo de la Escalera and Jos M Armingol, "Detection and Tracking of Pedestrians in Infrared Images", *International Conference on Signals, Circuits and Systems*, 2009

[5] Wai Kit Wong, Poi Ngee Tan, Chu Kiong Loo and Way Soong Lim, "An Effective Surveillance System Using Thermal Camera", *International Conference on Signal Acquisition and Processing*, 2009.

[6] www.imagefusion.org

[7] H.B.Mitchell, Image Fusion: Theories, Techniques and Applications, Springer, 2010.

[8] K. Han and G. N. DeSouza, "Instantaneous Geo-Location of Multiple Targets from Monocular Airborne Video," Proceedings of 2002 IEEE International Conference on Robotics and Automation, May 2002, Washington DC, USA.

[9] K. Han and G. N. DeSouza, "Multiple Targets Geolocation using SIFT and Stereo Vision on Airborne Video Sequences", The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems October 11-15, 2009 St. Louis, USA.

[10] D. B. Barber, J. Redding, T. W. McLain, R. W. Beard, and C. Taylor, "Vision-based target geo-location using a fixed-wing miniature air vehicle," Journal of Intelligent Robotics Systems, vol. 47, pp. 361–382.

[11] Redding, J. D., "Vision based target localization from a small fixed-wing unmanned air vehicle", Master's thesis, Brigham, Young University, Provo, Utah 84602.




DOI: 01.IJIT.01.02.33